WM-DIFF: WORLD MODEL-GUIDED PLANNING FOR ANGULAR DIFFUSION IN PROTEIN BACKBONE GENERATION

Anonymous authors
Paper under double-blind review

ABSTRACT

Internal-coordinate diffusion provides a translation and rotation-invariant representation for protein backbone generation, but existing samplers remain largely myopic by optimizing only the current denoising state without considering future structural consequences. Due to serial forward kinematics, small angular perturbations can accumulate into large Cartesian deviations and reduce sampling stability. We introduce WM-Diff, a World Model-guided planning framework that equips angular diffusion with lookahead capability through horizon-conditioned future prediction. WM-Diff learns future steric cost-to-go from deterministic reverse-diffusion rollouts of a frozen backbone generator rather than instantaneous structural penalties. During sampling, the World Model provides differentiable future-cost feedback to steer trajectories toward lower downstream steric cost. On an unconditional protein backbone generation benchmark, WM-Diff improves self-consistency designability from 27.5% to 52.0% while reducing steric clashes from 15.0 to 1.8 without post-hoc structural relaxation. Mechanistic analyses show that future-cost prediction provides a more effective planning signal than instantaneous objectives while preserving structural diversity. These results demonstrate the effectiveness of World Model-guided planning for protein backbone generation.

1 INTRODUCTION

De novo protein backbone generation provides structural scaffolds for subsequent sequence design, enabling downstream sequence optimization and functional protein design (Jumper et al., 2021; Yuan et al., 2025; Watson et al., 2023). Diffusion models have recently emerged as powerful generative frameworks for this task by modeling protein structures through different geometric representations (Ho et al., 2020; Song et al., 2020; Hoogeboom et al., 2022). FrameDiff models residue-level rigid frames and estimates $\mathrm{SE}(3)$-equivariant scores in Cartesian frame space (Yim et al., 2023), whereas FoldingDiff represents backbones as sequences of bond and dihedral angles and performs diffusion directly in internal-coordinate space (Wu et al., 2024). Internal coordinates remove dependence on global translation and rotation, allowing angular diffusion to operate without explicit global-pose modeling. WM-Diff builds on this angular formulation.

This representation nevertheless introduces geometric sensitivity during Cartesian reconstruction. Forward kinematics composes local transformations along the chain, so perturbing an upstream angle changes the positions of downstream atoms. The resulting lever-arm effect can amplify residual angular errors into larger Cartesian displacements and can create or worsen steric overlaps. It also makes direct World Model guidance difficult to scale: differentiating a steep coordinate-space penalty through the kinematic chain can produce poorly scaled angular gradients in sensitive configurations.

A direct guidance strategy evaluates a steric penalty on the backbone reconstructed from the current noisy state x_t. Such an objective can capture long-range contacts in the current structure, but it does not explicitly estimate the steric consequence after subsequent reverse-diffusion steps. This motivates a future-aware surrogate that predicts downstream steric cost without differentiating through a multi-step rollout at every sampling update.

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

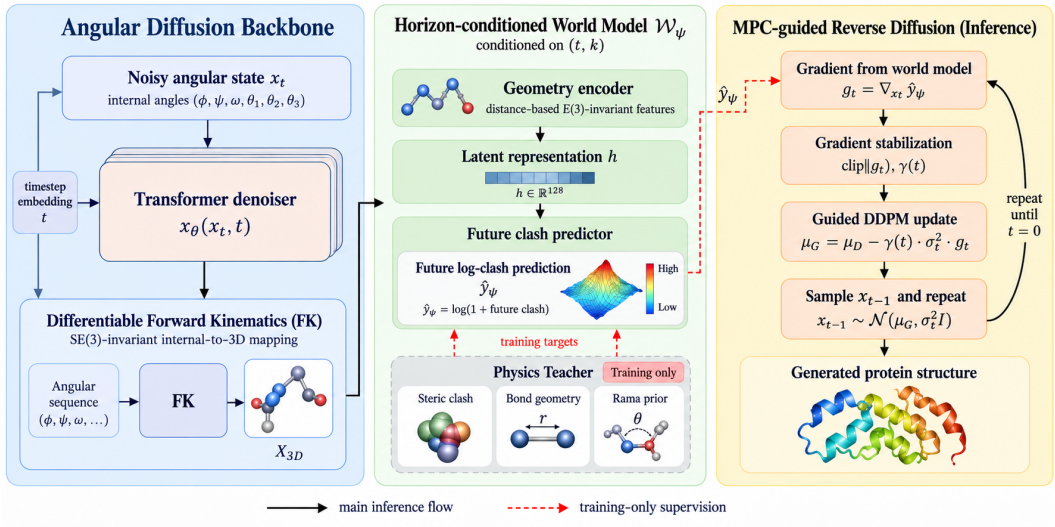


Figure 1: **WM-Diff framework.**

To address this issue, we formulate reverse angular diffusion as a lookahead guidance problem. Given a noisy angular state x_t, WM-Diff uses a horizon-conditioned predictor, termed the World Model $W_\psi(x_t, t, k)$, to estimate the log-transformed steric cost-to-go after k deterministic denoising steps performed by a frozen angular diffusion model. The frozen diffusion model defines the reverse rollout dynamics; W_ψ does not learn a separate state-transition head or predict a complete future protein structure.

During sampling, the gradient of the predicted future cost with respect to x_t shifts the reverse transition toward states with lower predicted downstream steric cost. This is related to guided diffusion, where gradients of auxiliary objectives modify reverse trajectories while retaining the pretrained generative model (Dhariwal & Nichol, 2021; Song et al., 2020). Because the learned surrogate does not remove the kinematic sensitivity of the angular representation, WM-Diff separately controls the applied guidance using gradient-norm clipping and a guidance scale that decreases linearly during denoising.

The resulting framework, WM-Diff(Figure 1), makes three contributions:

- **Receding-horizon guidance for angular diffusion.** We formulate reverse diffusion in internal-coordinate space as a receding-horizon guidance problem. This formulation explicitly accounts for the downstream steric consequences of angular updates under serial forward kinematics, extending guidance beyond instantaneous geometric objectives.
- **A horizon-conditioned steric cost-to-go model.** We develop an $\mathrm{E}(3)$-invariant World Model $W_\psi(x_t, t, k)$ that predicts the normalized log-transformed steric cost after a deterministic k-step reverse rollout. Its differentiable cost gradient is incorporated into the reverse transition through a time-dependent guidance schedule and global gradient-norm control, providing a computationally tractable lookahead mechanism for angular diffusion.
- **Controlled evaluation of lookahead guidance.** We evaluate WM-Diff on an unconditional single-chain backbone generation benchmark and perform ablations over rollout horizon, temporal scheduling, guidance strength, and gradient clipping. Under the evaluated protocol, WM-Diff improves self-consistency designability from 27.5% to 52.0% and reduces the mean severe clash count from 15.0 to 1.8 without post-hoc structural relaxation. The experiments further compare learned cost-to-go guidance with instantaneous analytical objectives and characterize the conditions under which lookahead guidance is effective.

2 Background and Motivation

Angular Representation and Kinematic Motivation. Following FoldingDiff, we represent a protein backbone as an angular sequence $x \in [-\pi, \pi)^M$, where $M = 6(N-1)$ includes three

bond-angle variables and three backbone torsion angles per residue transition (Wu et al., 2024). Internal coordinates provide invariance to rigid-body transformations, but the periodic geometry of angular space requires intrinsic distances. We therefore use the wrapped angular distance $d_{\mathbb{T}}(x, y) = \|\mathrm{wrap}(x - y)\|_2$ and a wrapped Gaussian diffusion process.

Differentiable Forward Kinematics and Lever-Arm Effect. Angular states are mapped to Cartesian backbone coordinates through a differentiable forward-kinematics operator $\mathbf{X} = \mathcal{FK}(x)$ based on NeRF (Parsons et al., 2005). Because the kinematic chain is serial, perturbations of upstream angular variables can influence distant downstream atoms. For an upstream torsion angle ϕ_j, the resulting Cartesian sensitivity follows the classical lever-arm relationship (Li et al., 1994; Mueller, 2019):

$$\frac{\partial \mathbf{X}_i}{\partial \phi_j} = \mathbf{u}_j \times (\mathbf{X}_i - \mathbf{p}_j). \tag{1}$$

which indicates that long-range geometric sensitivity can increase with chain distance. The detailed derivation is provided in Appendix A.1. This motivates a future-aware guidance strategy that predicts downstream structural consequences along the reverse trajectory rather than optimizing only instantaneous geometric penalties at the current noisy state(Rawlings et al., 2020).

3 Method: Model-Based Planning in Angular Diffusion

Internal-coordinate diffusion provides rigid-body invariance and compact geometric parameterization, but the reverse denoising process does not explicitly consider the downstream structural consequences of current angular updates. To address this limitation, we formulate guided sampling as a model-based planning problem over the diffusion trajectory(Rawlings et al., 2020; Hafner et al., 2019).

WM-Diff augments a pretrained angular diffusion model with a horizon-conditioned World Model that predicts future steric cost-to-go(Hafner et al., 2019; Schulman et al., 2015). During reverse diffusion, the predicted future cost is used to provide differentiable lookahead guidance, encouraging intermediate states toward trajectories with improved downstream structural feasibility.

The diffusion backbone remains frozen after training and defines the base reverse rollout dynamics. The World Model does not replace the generative model; instead, it acts as a lightweight predictive model that estimates future structural consequences under the existing diffusion trajectory distribution.

3.1 Horizon-Conditioned World Model as a Steric Cost-to-Go Predictor

Steric cost target. A central challenge of angular-space guidance is that direct optimization of instantaneous Cartesian objectives can produce unstable gradients due to the nonlinear relationship between angular variables and Cartesian coordinates(Dhariwal & Nichol, 2021). Instead of using the current steric state alone, we train the World Model to predict future steric feasibility after several reverse diffusion steps.

We supervise the World Model using a continuous steric-overlap objective:

$$E_{\mathrm{clash}}(\mathbf{X}) = \sum_{(i,j)\in\mathcal{P}_{\mathrm{nb}}} \left[\max\left(r_{\mathrm{vdw},ij} - \|\mathbf{X}_i - \mathbf{X}_j\|_2, 0\right)\right]^2. \tag{2}$$

where $\mathcal{P}_{\mathrm{nb}}$ denotes the set of valid non-bonded heavy-atom pairs. The operator $\max(\cdot, 0)$ ensures that only positive van der Waals overlaps contribute to the loss. Thus, pairs whose distance exceeds the sum of their van der Waals radii receive zero penalty rather than an artificial attractive force. This objective is used as a differentiable steric-overlap surrogate for training the World Model and is not intended to represent a complete molecular energy function.

Horizon-conditioned rollout. Let θ denote the frozen diffusion parameters. Given a noisy angular state x_t, the pretrained diffusion model defines a reverse trajectory(Song et al., 2020). We construct future states through deterministic k-step rollouts:

$$\hat{x}_{t-m-1} = f_\theta(\hat{x}_{t-m}, t-m), \qquad m = 0, \ldots, k-1. \tag{3}$$

This rollout formulation allows the World Model to evaluate the consequences of current denoising decisions over a future horizon, rather than relying only on instantaneous geometric information.

The terminal rollout state is used to compute the future steric target:

$$y_{t,k} = \log\left(1 + \frac{E_{\text{clash}}(\mathcal{FK}(\hat{x}_{t-k}))}{E_{\text{ref}}}\right). \tag{4}$$

where E_{ref} provides normalization.

The horizon-conditioned World Model predicts this future cost:

$$\hat{y}_\psi(x_t, t, k) = W_\psi(x_t, t, k). \tag{5}$$

and is optimized using:

$$\mathcal{L}_{\text{world}} = \mathbb{E}\left[(W_\psi(x_t, t, k) - y_{t,k})^2\right]. \tag{6}$$

Therefore, W_ψ learns an amortized approximation of future steric cost-to-go conditioned on the current diffusion state and planning horizon(Sutton et al., 1998).

Invariant geometric representation. At each diffusion step, angular states are converted into Cartesian coordinates through differentiable forward kinematics. The World Model receives geometric information through invariant pairwise distance features computed from reconstructed C_α coordinates:

$$D_{ij}^{(t)} = \|\mathbf{C}_{t,i} - \mathbf{C}_{t,j}\|_2. \tag{7}$$

These geometric features are combined with timestep and horizon embeddings to predict future steric cost. Since pairwise distances are preserved under global rigid transformations, the representation is invariant to translation, rotation, and reflection(Thomas et al., 2018).

3.2 Receding-Horizon World Model Guidance

During inference, the World Model is frozen and used as a differentiable future-cost surrogate(Hafner et al., 2019; Rawlings et al., 2020). Rather than optimizing an instantaneous physical penalty, WM-Diff biases the reverse diffusion trajectory toward states with lower predicted downstream steric cost(Dhariwal & Nichol, 2021; Ho & Salimans, 2022):

$$\pi_{t,k}(x) \propto p_t(x) \exp\left(-\gamma(t) W_\psi(x, t, k)\right). \tag{8}$$

The corresponding guidance direction is obtained from the gradient of the predicted future cost(Dhariwal & Nichol, 2021):

$$g_t^{\text{raw}} = \nabla_{x_t} W_\psi(x_t, t, k). \tag{9}$$

This formulation enables the sampler to incorporate future structural consequences into each denoising step. Similar to model predictive control, the future cost is repeatedly evaluated after state updates, forming a receding-horizon planning procedure over the reverse diffusion trajectory(Rawlings et al., 2020).

Using a local Gaussian guidance approximation, the reverse transition is modified as:

$$p_{\text{guided}}(x_{t-1}|x_t) \approx \mathcal{N}_{\text{wrap}}\left(x_{t-1}; \mu_\theta^{\text{base}} - \gamma(t)\sigma_t^2 \tilde{g}_t, \sigma_t^2 I\right). \tag{10}$$

3.3 KINEMATIC GRADIENT STABILIZATION AND SAMPLING

Because angular perturbations can produce amplified downstream effects through forward kinematics, the guidance magnitude is stabilized using global gradient-norm clipping(Pascanu et al., 2013):

$$\tilde{g}_t = g_t^{\mathrm{raw}} \min\left(1, \frac{\tau}{\max(\|g_t^{\mathrm{raw}}\|_2, \epsilon_g)}\right). \tag{11}$$

When $\|g_t^{\mathrm{raw}}\|_2 > \tau$, this operation becomes

$$\tilde{g}_t = \tau \frac{g_t^{\mathrm{raw}}}{\|g_t^{\mathrm{raw}}\|_2}. \tag{12}$$

so the update preserves the direction of the World Model gradient while replacing its original magnitude with the fixed norm τ. Therefore, in regimes where clipping is active for most denoising steps, the method should be interpreted as approximately normalized gradient guidance rather than as direct use of the calibrated absolute cost magnitude. The guidance scale $\gamma(t)$ and the threshold τ still determine the resulting update size.

The final guided reverse update is:

$$x_{t-1} = \mathrm{wrap}\left(\mu_\theta^{\mathrm{base}} - \gamma(t)\sigma_t^2 \tilde{g}_t + \sigma_t z_t\right). \tag{13}$$

All angular states are wrapped after each reverse update. When the selected horizon is unavailable, the sampler follows the corresponding unguided diffusion transition.

4 EXPERIMENTS

The main experiments emphasize the evidence required to validate the core contributions of WM-Diff. Additional visualization results and extended sensitivity tables are provided in the appendix.

Experimental organization. The main experiments focus on the evaluations required to establish the effectiveness of WM-Diff, including benchmark comparison, guidance effect analysis, and core component analysis. Additional robustness studies, random-seed evaluations, and extended sensitivity analyses are provided in the appendix to avoid interrupting the main narrative.

In this section, we evaluate WM-Diff from three perspectives. First, we assess its generation quality on an unconditional single-chain protein-backbone benchmark using self-consistency evaluation. Second, we perform ablation studies to quantify the contribution of individual components, including guidance strength and gradient stabilization. Finally, we conduct mechanistic analyses of the lookahead guidance signal and its comparison with instantaneous analytical objectives.

4.1 EXPERIMENTAL SETUP

Datasets and Training Protocol. All models are trained on the non-redundant CATH domain dataset (Orengo et al., 1997) using a 1000-step cosine angular noise schedule in internal-coordinate space.

The diffusion backbone adopts a 12-layer BERT-style Transformer architecture(Vaswani et al., 2017; Devlin et al., 2019). For fair comparison, WM-Diff and all baseline models are evaluated under consistent optimization and sampling settings, including training schedules, batch configurations, and evaluation protocols.

Evaluation Protocol. For each generated unrelaxed backbone, we evaluate self-consistency through a ProteinMPNN–ESMFold pipeline (Dauparas et al., 2022; Lin et al., 2023). ProteinMPNN samples $R = 8$ candidate sequences per backbone at temperature $T = 0.1$, and ESMFold predicts the corresponding structures. The best TM-score among the eight predicted structures is used as the self-consistency score (scTM). Following previous evaluation protocols(Watson et al., 2023; Yim et al., 2023), computational designability is reported as the fraction of generated backbones with $\mathrm{scTM} \geq 0.5$.

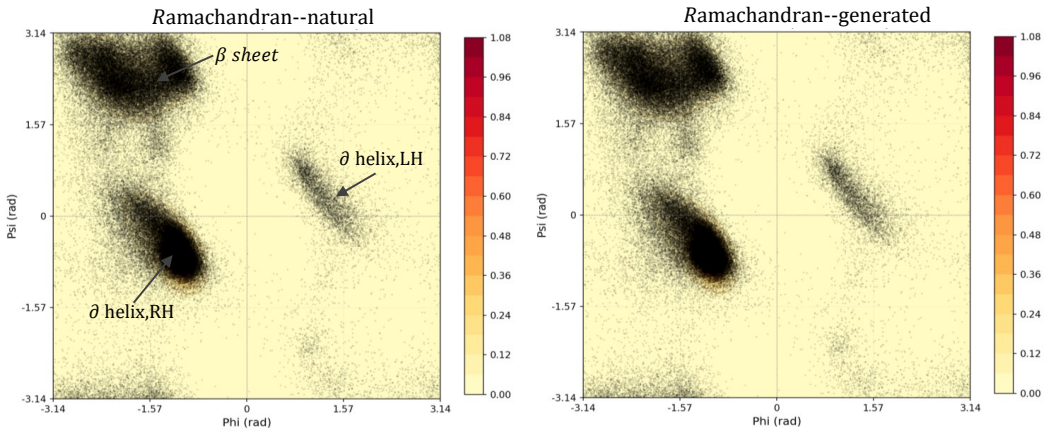


Figure 2: **Recovery of natural backbone conformational preferences.** Ramachandran density distributions of backbone dihedral angles. Joint distributions of ϕ and ψ angles comparing natural CATH domains (*left*) with unconditional backbones generated by WM-Diff (*right*). Arrows indicate characteristic density regions corresponding to right-handed α-helices, left-handed α-helices, and β-sheet conformations.

We further evaluate structural quality using Clash Count, defined as the number of valid non-bonded heavy-atom pairs with van der Waals overlap larger than 0.4 Å on unrelaxed coordinates. This metric is distinct from the continuous steric objective E_{clash} used during training. DSSP-assigned β-sheet content is reported as a descriptive structural statistic, while within-length-bin pairwise TM-score is used to measure structural similarity and diversity(Kabsch & Sander, 1983).

The benchmark contains 782 generated backbones under the reported configuration. Detailed masking rules, pair exclusions, and complete metric definitions are provided in Appendix B.

4.2 Main Results: Unconditional Generation

Recovery of Natural Backbone Geometry. We first evaluate whether WM-Diff recovers realistic backbone conformational preferences by examining the distribution of backbone dihedral angles.

Figure 2 compares the Ramachandran density distribution(Ramachandran, 1963; Carugo & Djinović-Carugo, 2013) of natural CATH domains with unconditional backbones generated by WM-Diff.

The generated structures recover the major conformational basins associated with right-handed α-helices, left-handed α-helices, and β-sheet regions(Cintas, 2002). Although this marginal-angle analysis does not fully characterize global stereochemical validity, it demonstrates that WM-Diff preserves common local backbone geometries when generating structures directly in internal-coordinate space.

Self-Consistency Folding Evaluation. To assess whether generated backbones can correspond to compatible amino-acid sequences and folded structures, we evaluate self-consistency using an inverse-folding and structure prediction pipeline.

As illustrated in Figure 3(a), each generated backbone is first inverse-folded into an amino-acid sequence using ProteinMPNN, followed by structure prediction using ESMFold. The structural agreement between the generated backbone and the predicted structure is measured by the self-consistency TM-score (scTM)(Yim et al., 2023; Zhang & Skolnick, 2005).

Figure 3(b) presents the resulting scTM distribution across 782 generated backbones.

Quantitative Designability Evaluation. Table 1 summarizes the unconditional generation performance of WM-Diff against representative protein backbone generative baselines, including FoldingDiff (Wu et al., 2024), ProtDiff (Trippe et al., 2022), and ProteinSGM (Lee et al., 2023).

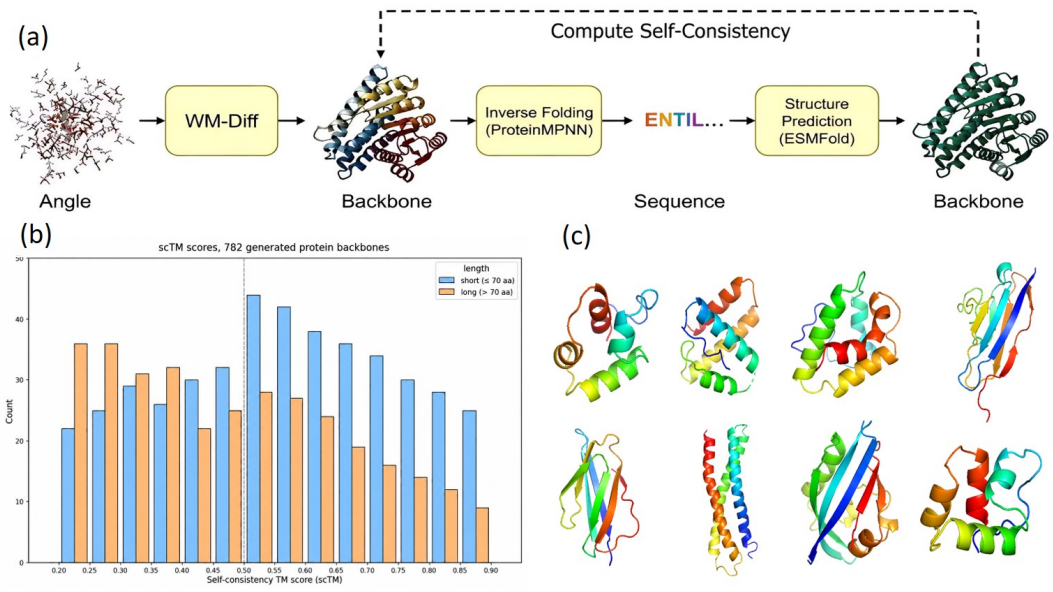


Figure 3: **Unconditional generation and self-consistency designability benchmark. (a)** Self-consistency designability pipeline: generated backbones are inverse-folded using ProteinMPNN, followed by ESMFold structure prediction. **(b)** Distribution of scTM scores across 782 generated backbones. **(c)** Representative generated backbones across different chain lengths.

Compared with unguided WM-Diff, World-Model-guided sampling consistently improves self-consistency designability while substantially reducing severe steric clashes. In particular, with a prediction horizon of $k = 4$, WM-Diff achieves a designability rate of 52.0%, a mean scTM of 0.59, and an average steric clash count of 1.8, without post-hoc structural relaxation.

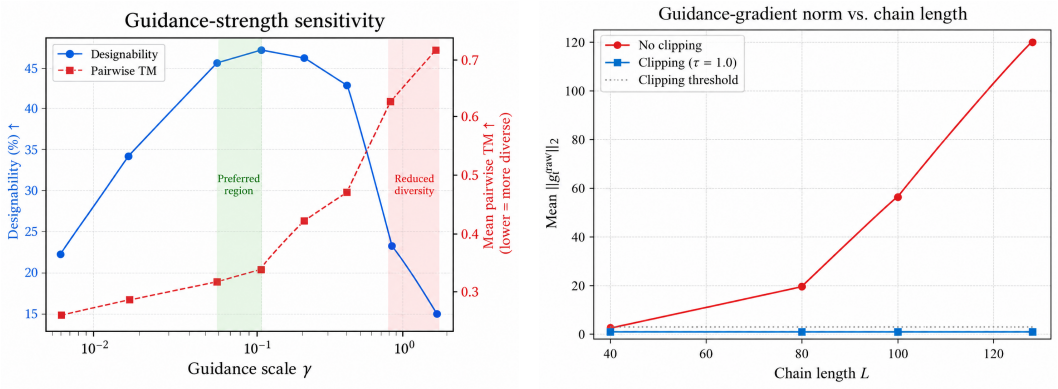
These results are consistent with the intended role of the learned World Model, which provides an auxiliary estimate of downstream structural feasibility during reverse diffusion. By incorporating predicted future evolution, the guidance signal can steer sampling away from locally plausible but downstream unfavorable conformations. However, the present evaluation is limited to unconditional single-chain generation and does not establish whether the same mechanism generalizes to conditional generation, motif scaffolding, multi-chain generation, or function-constrained design.

World-Model guidance also shifts the secondary-structure distribution. The guided $k = 4$ configuration contains 32.0% DSSP-assigned β-sheet residues, compared with 12.0% for unguided WM-Diff. This 20-percentage-point increase reflects a substantial change in secondary-structure composition. However, β-sheet content alone does not measure sheet topology, strand pairing, packing quality, or hydrogen-bond organization, and should therefore not be interpreted as direct evidence of improved β-sheet structural quality.

Table 1: **Unconditional backbone generation benchmark under internal-coordinate diffusion.** Results on 782 generated backbones without post-hoc relaxation. Higher scTM and lower Clash Count indicate better performance.

Method	**Designable (%)** ↑	**Mean scTM** ↑	**Clash Count** ↓	**β-sheet (%)**
FoldingDiff	22.5%	0.38	20.0	8.0%
ProtDiff	12.5%	0.39	10.5	9.0%
ProteinSGM	44.4%	0.44	5.1	12.1%
WM-Diff (Base, w/o guidance)	27.5%	0.42	15.0	12.0%
WM-Diff (Guided, Horizon $k = 1$)	45.0%	0.53	4.5	25.0%
WM-Diff (Guided, Horizon $k = 4$)	**52.0%**	**0.59**	**1.8**	**32.0%**

Representative Structural Examples. Finally, Figure 3(c) presents representative WM-Diff generated backbones spanning different chain lengths. These examples demonstrate that WM-Diff produces diverse backbone conformations while maintaining strong self-consistency with their corresponding ESMFold predictions, highlighting the ability of World Model-guided sampling to explore diverse structural conformational spaces.



(a) (b)

Figure 4: **Guidance ablations. (a)** Guidance strength controls the trade-off between designability and diversity, with moderate guidance achieving the best balance. **(b)** Gradient clipping stabilizes World Model guidance by preventing length-dependent gradient growth.

4.3 Ablation Studies

4.3.1 Effect of Guidance Strength

Guidance-strength sensitivity. We evaluate how the magnitude of World Model guidance affects generation behavior by sweeping the guidance coefficient γ.

This experiment is designed as a sensitivity analysis of the guidance mechanism while retaining the benchmark-consistent evaluation protocol. The unguided ($\gamma = 0$) and selected guidance ($\gamma = 0.10$) settings are aligned with the corresponding configurations reported in Table 1; intermediate and larger guidance values are evaluated as sensitivity points.

As shown in Figure 4a, removing World Model guidance ($\gamma = 0$) corresponds to the unguided benchmark setting and achieves 27.5% designability. Introducing moderate guidance improves designability. The setting $\gamma = 0.10$ achieves the best observed trade-off in this sweep, reaching 52.0% designability while maintaining structural diversity. Increasing the guidance magnitude beyond this regime reduces the benefit of guidance. For $\gamma \geq 0.80$, generated structures become increasingly similar, as reflected by the increase in mean pairwise TM-score and reduction of structural diversity.

The final benchmark configuration reported in Table 1 combines the selected guidance setting with the complete sampling configuration, including the chosen sampling horizon and other stabilization components. Therefore, the benchmark designability and the guidance-sweep designability values are not directly interchangeable.

4.3.2 Effect of Gradient Stabilization

Gradient Instability in Long-chain Generation. The analysis in Section 2. shows that forward kinematics can amplify angular perturbations through the lever-arm effect. Because World Model guidance is differentiated with respect to the current angular state, this kinematic sensitivity can lead to large guidance gradients for long protein chains. We therefore evaluate the effect of gradient clipping on guidance stability across different chain lengths.

Empirical Gradient Amplification. Figure 4b reports the mean guidance gradient norm for chain lengths $L \in \{40, 80, 100, 128\}$. Without gradient clipping, the measured gradient norm increases substantially with chain length, from 2.5 at $L = 40$ to 18.3 at $L = 80$, 55.0 at $L = 100$, and 120.0 at $L = 128$. This amplification is accompanied by unstable reverse diffusion sampling for long chains. At $L = 128$, the corresponding self-consistency designability decreases from 55.0% to 10.0% under the unclipped setting in the long-chain stress-test configuration. This analysis is performed on the long-chain subset and is not intended to reproduce the full benchmark distribution reported in Table 1.

Effect of Gradient Clipping. Applying the gradient-norm clipping strategy defined in Section 3 bounds the magnitude of the World Model guidance gradient during sampling. As shown in Fig-

Table 2: **Mechanistic analysis of World Model guidance.**

Schedule	Form	Designability (%) ↑
Linear Decay	$\gamma_0(t/T)$	**52.0**
Constant	γ_0	41.0
Early	$\gamma_0\mathbb{I}(t > 0.5T)$	46.0
Late	$\gamma_0\mathbb{I}(t \leq 0.5T)$	32.0

(a) Temporal guidance schedules.

Guidance	Designability (%) ↑	Clash ↓
World Model	**52.0**	1.8
Analytical	35.0	1.8
Clash	28.0	**0.9**
Clash+Bond	32.0	1.4
All terms	36.0	1.9

(b) Learned vs. analytical guidance.

ure 4b, using $\tau = 1.0$ limits the gradient norm across the evaluated chain lengths. For $L = 128$, clipped guidance maintains a designability rate of 42.0%, compared with 10.0% in the corresponding unclipped long-chain stress-test setting.

These results provide empirical support for gradient clipping as an important stabilization component of WM-Diff, particularly for longer protein backbones.

4.4 Mechanistic Analysis

We further analyze two aspects of World Model guidance: (i) how guidance should be distributed across the reverse trajectory, and (ii) whether learned future-cost guidance provides advantages over instantaneous analytical objectives.

Temporal Dynamics of Guidance. To examine how guidance timing affects sampling, we compare several temporal schedules while keeping the remaining guided-sampling configuration fixed. The Linear Decay schedule achieves the highest designability (52.0%), followed by Early-focused (46.0%), Constant (41.0%), and Late-focused (32.0%) guidance.

The stronger performance of early-stage guidance suggests that applying World Model gradients during high-noise reverse diffusion stages can better influence coarse structural organization before local refinement. However, this ablation does not establish a strict causal separation between topology formation and local geometry refinement.

Learned World Model vs. Analytical Energy Guidance. We further examine whether the learned World Model can be replaced by guidance from analytical geometric objectives. As shown in Table 2, all analytical variants achieve substantially lower designability than the learned World Model. In particular, clash-only guidance reduces the average clash count to 0.9, but achieves only 28.0% designability, compared with 52.0% designability and 1.8 clashes for the World Model.

These results support the advantage of horizon-conditioned future-cost prediction over instantaneous analytical penalties. They do not imply that the World Model represents a complete molecular energy landscape or molecular dynamics model.

5 Conclusion and Discussion

We introduced WM-Diff, a World Model-guided planning framework that augments internal-coordinate diffusion with future-aware steric guidance for protein backbone generation. By learning a horizon-conditioned steric cost-to-go surrogate and leveraging its gradients during reverse diffusion, WM-Diff enables the sampler to anticipate downstream geometric consequences of angular updates without requiring post-hoc structural relaxation. Experiments on unconditional single-chain backbone generation demonstrate substantial improvements in self-consistency designability and a strong reduction in severe steric clashes, while ablation studies reveal the importance of planning horizon, guidance scheduling, guidance strength, and gradient regulation. Mechanistic analyses further show that predicting future steric cost provides a more effective guidance signal than instantaneous objectives while maintaining structural diversity. The current World Model remains a residue-level learned steric surrogate rather than a complete molecular energy or dynamics model, and the current evaluation is limited to unconditional single-chain backbone generation. Future work will investigate extensions toward atomistic and multi-chain representations, conditional and motif-constrained generation, improved guidance calibration, and experimental validation.

AI USE STATEMENT

Generative AI tools were used solely to assist with language polishing, grammatical refinement, and improving manuscript clarity and readability. All research ideas, methodological design, experimental procedures, results, analyses, and scientific conclusions were developed and verified by the authors. No AI tools generated experimental results or made scientific decisions.

REFERENCES

Oliviero Carugo and Kristina Djinović-Carugo. Half a century of ramachandran plots. *Biological Crystallography*, 69(8):1333–1341, 2013.

Pedro Cintas. Chirality of living systems: a helping hand from crystals and oligopeptides. *Angewandte Chemie International Edition*, 41(7):1139–1145, 2002.

Justas Dauparas, Ivan Anishchenko, Nathaniel Bennett, Hua Bai, Robert J Ragotte, Lukas F Milles, Basile IM Wicky, Alexis Courbet, Rob J de Haas, Neville Bethel, et al. Robust deep learning–based protein sequence design using proteinmpnn. *Science*, 378(6615):49–56, 2022.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pp. 4171–4186, 2019.

Prafulla Dhariwal and Alexander Nichol. Diffusion models beat gans on image synthesis. *Advances in neural information processing systems*, 34:8780–8794, 2021.

Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pp. 1321–1330. PMLR, 2017.

Danijar Hafner, Timothy Lillicrap, Jimmy Ba, and Mohammad Norouzi. Dream to control: Learning behaviors by latent imagination. *arXiv preprint arXiv:1912.01603*, 2019.

Jonathan Ho and Tim Salimans. Classifier-free diffusion guidance. *arXiv preprint arXiv:2207.12598*, 2022.

Jonathan Ho, Ajay Jain, and Pieter Abbeel. Denoising diffusion probabilistic models. *Advances in neural information processing systems*, 33:6840–6851, 2020.

Emiel Hoogeboom, Vıctor Garcia Satorras, Clément Vignac, and Max Welling. Equivariant diffusion for molecule generation in 3d. In *International conference on machine learning*, pp. 8867–8887. PMLR, 2022.

John Jumper, Richard Evans, Alexander Pritzel, Tim Green, Michael Figurnov, Olaf Ronneberger, Kathryn Tunyasuvunakool, Russ Bates, Augustin Žídek, Anna Potapenko, et al. Highly accurate protein structure prediction with alphafold. *nature*, 596(7873):583–589, 2021.

Wolfgang Kabsch and Christian Sander. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers: Original Research on Biomolecules*, 22(12):2577–2637, 1983.

Jin Sub Lee, Jisun Kim, and Philip M Kim. Score-based generative modeling for de novo protein design. *Nature Computational Science*, 3(5):382–392, 2023.

Zexiang Li, Shankar Sastry, and R Murray. *A mathematical introduction to robotic manipulation*. CRC press Boca Raton, FL, 1994.

Zeming Lin, Halil Akin, Roshan Rao, Brian Hie, Zhongkai Zhu, Wenting Lu, Nikita Smetanin, Robert Verkuil, Ori Kabeli, Yaniv Shmueli, et al. Evolutionary-scale prediction of atomic-level protein structure with a language model. *Science*, 379(6637):1123–1130, 2023.

Andreas Mueller. Modern robotics: Mechanics, planning, and control [bookshelf]. *IEEE Control Systems Magazine*, 39(6):100–102, 2019.

Christine A Orengo, Alex D Michie, Susan Jones, David T Jones, Mark B Swindells, and Janet M Thornton. Cath–a hierarchic classification of protein domain structures. *Structure*, 5(8):1093–1109, 1997.

Jerod Parsons, J Bradley Holmes, J Maurice Rojas, Jerry Tsai, and Charlie EM Strauss. Practical conversion from torsion space to cartesian space for in silico protein synthesis. *Journal of computational chemistry*, 26(10):1063–1068, 2005.

Razvan Pascanu, Tomas Mikolov, and Yoshua Bengio. On the difficulty of training recurrent neural networks. In *International conference on machine learning*, pp. 1310–1318. Pmlr, 2013.

Gopalasamudram Narayana Ramachandran. Stereochemistry of polypeptide chain configurations. *J. Mol. Biol.*, 7:95–99, 1963.

James B Rawlings, David Q Mayne, and Moritz M Diehl. Model predictive control: theory, computation, and design. *(No Title)*, 2020.

John Schulman, Sergey Levine, Pieter Abbeel, Michael Jordan, and Philipp Moritz. Trust region policy optimization. In *International conference on machine learning*, pp. 1889–1897. Pmlr, 2015.

Yang Song, Jascha Sohl-Dickstein, Diederik P Kingma, Abhishek Kumar, Stefano Ermon, and Ben Poole. Score-based generative modeling through stochastic differential equations. *arXiv preprint arXiv:2011.13456*, 2020.

Charles Spearman. The proof and measurement of association between two things. 1961.

Richard S Sutton, Andrew G Barto, and Andrew Barto. *Reinforcement learning: An introduction*, volume 1. MIT press Cambridge, 1998.

Nathaniel Thomas, Tess Smidt, Steven Kearnes, Lusann Yang, Li Li, Kai Kohlhoff, and Patrick Riley. Tensor field networks: Rotation-and translation-equivariant neural networks for 3d point clouds. *arXiv preprint arXiv:1802.08219*, 2018.

Brian L Trippe, Jason Yim, Doug Tischer, David Baker, Tamara Broderick, Regina Barzilay, and Tommi Jaakkola. Diffusion probabilistic modeling of protein backbones in 3d for the motif-scaffolding problem. *arXiv preprint arXiv:2206.04119*, 2022.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

Joseph L Watson, David Juergens, Nathaniel R Bennett, Brian L Trippe, Jason Yim, Helen E Eisenach, Woody Ahern, Andrew J Borst, Robert J Ragotte, Lukas F Milles, et al. De novo design of protein structure and function with rfdiffusion. *Nature*, 620(7976):1089–1100, 2023.

Cort J Willmott and Kenji Matsuura. Advantages of the mean absolute error (mae) over the root mean square error (rmse) in assessing average model performance. *Climate research*, 30:79–82, 2005.

Kevin E Wu, Kevin K Yang, Rianne Van den Berg, Sarah Alamdari, James Y Zou, Alex X Lu, and Ava P Amini. Protein structure generation via folding diffusion. *Nature communications*, 15(1): 1059, 2024.

Jason Yim, Brian L Trippe, Valentin De Bortoli, Emile Mathieu, Arnaud Doucet, Regina Barzilay, and Tommi Jaakkola. Se (3) diffusion model with application to protein backbone generation. *arXiv preprint arXiv:2302.02277*, 2023.

Chaohao Yuan, Songyou Li, Geyan Ye, Yikun Zhang, Long-Kai Huang, Wenbing Huang, Wei Liu, Jianhua Yao, and Yu Rong. Annotation-guided protein design with multi-level domain alignment. In *Proceedings of the 31st ACM SIGKDD Conference on Knowledge Discovery and Data Mining V. 1*, pp. 1855–1866, 2025.

Yang Zhang and Jeffrey Skolnick. Tm-align: a protein structure alignment algorithm based on the tm-score. *Nucleic acids research*, 33(7):2302–2309, 2005.

A Implementation Details and Reproducibility

This appendix provides implementation details for reproducing WM-Diff. The mathematical formulation of the diffusion process, World Model objective, guided reverse transition, and gradient stabilization strategy are described in Section 3. We first introduce the forward-kinematics transformation convention, followed by model architectures, optimization configurations, and sampling details.

A.1 Forward-Kinematics Transformation

We reconstruct Cartesian backbone coordinates from internal coordinates using the standard NeRF recursion. Let $\mathbf{X}_{i-3}$, $\mathbf{X}_{i-2}$, and $\mathbf{X}_{i-1}$ denote the three previously constructed atoms. For the next atom $\mathbf{X}_i$, let b_i be the bond length, θ_i the bond angle, and ϕ_i the signed dihedral angle.

Define the local orthonormal frame

$$\mathbf{e}_1 = \frac{\mathbf{X}_{i-1} - \mathbf{X}_{i-2}}{\|\mathbf{X}_{i-1} - \mathbf{X}_{i-2}\|_2}, \tag{14}$$

$$\mathbf{n} = \frac{(\mathbf{X}_{i-2} - \mathbf{X}_{i-3}) \times \mathbf{e}_1}{\|(\mathbf{X}_{i-2} - \mathbf{X}_{i-3}) \times \mathbf{e}_1\|_2}, \qquad \mathbf{e}_2 = \mathbf{n} \times \mathbf{e}_1. \tag{15}$$

The next Cartesian coordinate is then given by

$$\mathbf{X}_i = \mathbf{X}_{i-1} + b_i \left[-\cos(\theta_i)\mathbf{e}_1 + \sin(\theta_i) \left(\cos(\phi_i)\mathbf{e}_2 + \sin(\phi_i)\mathbf{n} \right) \right]. \tag{16}$$

The sign convention for ϕ_i follows the right-hand rule around $\mathbf{e}_1$(Li et al., 1994; Mueller, 2019).. The initial three atoms are placed using the fixed reference geometry, after which Eq. equation 16 is applied recursively. This construction preserves the prescribed bond lengths and bond angles while allowing the dihedral angle to rotate the downstream chain around the corresponding bond axis.

A.2 Diffusion Backbone Configuration

The diffusion backbone operates on the six-dimensional internal-coordinate representation introduced in Section 2, following the angular diffusion formulation of FoldingDiff (Wu et al., 2024), where each residue transition is represented by three bond-angle variables and three torsion-angle variables.

The backbone predicts the reverse diffusion transition parameters used during sampling. The architecture and diffusion hyperparameters are summarized in Table 3.

Table 3: Configuration of the angular diffusion backbone.

Component	Configuration
Architecture	BERT-style Transformer
Number of layers	12
Hidden dimension	384
Intermediate dimension	768
Attention heads	12
Maximum sequence length	128
Diffusion steps	1000
Noise schedule	Cosine schedule
Representation	Internal angular coordinates

All experiments use the same trained diffusion backbone when comparing different guidance strategies.

Table 4: Architecture configuration of the horizon-conditioned World Model.

Component	Configuration
Geometric input	Pairwise C_α distances
Distance encoder	DistGram
Number of RBF bases	32
Maximum distance	40.0 Å
Transformer layers	3
Hidden dimension	128
Attention heads	4
Dropout	0.1
Pooling	Masked mean pooling
Conditioning	Timestep and horizon embeddings
Output	Scalar log-transformed steric cost-to-go

A.3 World Model Architecture

The World Model $W_\psi(x_t, t, k)$ is implemented as a horizon-conditioned Transformer encoder that predicts future steric cost-to-go from noisy intermediate states, following the predictive modeling principle of world models (Hafner et al., 2019).

The model receives noisy angular states x_t, diffusion timestep embeddings t, and planning horizon embeddings k. To capture geometry-aware structural dependencies, the angular states are first converted into invariant geometric representations before being processed by the Transformer encoder.

Specifically, the model uses pairwise C_α distance features:

$$D_{ij}^{(t)} = \|\mathbf{C}_{\alpha,i}^{(t)} - \mathbf{C}_{\alpha,j}^{(t)}\|_2. \tag{17}$$

where Cartesian coordinates are reconstructed from angular states using the differentiable forward-kinematics mapping introduced in Section 2. The World Model architecture is summarized in Table 4.

The World Model predicts the future steric cost-to-go target defined in Eq. (4) and does not introduce additional dynamics prediction heads.

A.4 Training Configuration

The World Model is trained using detached deterministic rollouts generated by the frozen diffusion backbone. The optimization objective and loss function follow those defined in Section 3.

Training hyperparameters are summarized in Table 5.

Table 5: World Model training configuration.

Parameter	Value
Dataset	CATH v4.3
Minimum sequence length	40
Maximum sequence length	128
Optimizer	AdamW
Batch size	128
Weight decay	0.01
Optimizer gradient clipping	1.0
World Model objective	Mean squared prediction error
Planning horizons	$k \in \{1, 4\}$
Diffusion backbone during World Model training	Frozen

The diffusion backbone parameters remain frozen during World Model training, and optimization is performed only for the future steric cost prediction objective.

A.5 Sampling Configuration

During inference, WM-Diff applies the guided reverse transition defined in Eq. (10). Unless otherwise specified, all experiments use the default sampling configuration summarized in Table 6.

Table 6: Default WM-Diff sampling configuration used for the main benchmark.

Parameter	Value
Planning horizon	$k = 4$
Initial guidance scale	$\gamma_0 = 0.10$
Guidance schedule	Linear Decay
Gradient clipping threshold	$\tau = 1.0$
Main benchmark size	782 backbones
Seed-stability evaluation	5 independent sampling seeds

The guidance magnitude is reduced during reverse diffusion according to the schedule defined in Section 3. The gradient clipping operation is applied to the World Model guidance gradient before modifying the reverse transition.

A.6 Computational Details

All experiments are implemented in PyTorch. The diffusion backbone and World Model are trained in separate stages.

During guided inference, both models are kept frozen with respect to their parameters. Gradients are computed only with respect to the current angular state x_t to obtain the World Model guidance signal.

Training-time optimizer gradient clipping is distinct from the World Model guidance-gradient clipping applied during sampling.

B Data Processing and Evaluation Protocols

This appendix provides additional details on dataset preprocessing, generation settings, and evaluation metrics. Unless otherwise stated, all methods are evaluated using the same downstream evaluation pipeline without post-hoc structural relaxation or gradient-based structural cleanup.

B.1 Dataset Preprocessing

All diffusion models are trained on the non-redundant CATH v4.3 domain dataset(Orengo et al., 1997). Following the internal-coordinate representation introduced in Section 3, a protein backbone of length N is represented using six angular variables for each residue transition, including three bond angles and three backbone torsion angles. The resulting representation has dimension

$$M = 6(N - 1). \tag{18}$$

where each angular variable is wrapped into the periodic interval $[-\pi, \pi)$. This representation removes the need to explicitly model global translation and rotation.

The maximum sequence length used by the diffusion backbone is 128 residues. Shorter sequences are padded for minibatch training, and corresponding padding masks are applied during Transformer computation and feature aggregation. The same representation and masking strategy are used for both the unguided diffusion baseline and WM-Diff.

The forward diffusion process uses $T = 1000$ timesteps with a cosine noise schedule. During World Model evaluation, angular states are converted into Cartesian coordinates using the differentiable forward-kinematics operator FK. The reconstructed coordinates are then used to construct geometric features and compute the continuous clash penalty defined in Section 3.

B.2 Evaluation Pipeline

We evaluate unconditional backbone generation using a self-consistency evaluation pipeline. For each generated backbone $\mathbf{X}^{(a)}$, ProteinMPNNdauparas2022robust generates $R = 8$ candidate amino-acid sequences using sampling temperature $T = 0.1$. The same candidate number, sampling temperature, and random seed protocol are used for all compared methods. Each candidate sequence is subsequently folded independently using ESMFold.

Let $\widehat{\mathbf{X}}^{(a,r)}$ denote the ESMFold-predicted structure of the r-th candidate sequence generated for backbone $\mathbf{X}^{(a)}$. A generated backbone is considered designable when $\mathrm{scTM} \geq 0.5$. All methods use identical sequence generation, structure prediction, and alignment procedures to ensure that differences arise from the generated backbones rather than the downstream evaluation pipeline.

The main benchmark consists of 782 generated backbones. For length-stratified diversity analysis, structures are divided into short chains (≤ 70 residues) and long chains (> 70 residues), following the partition used in the main text. Pairwise structural similarity is computed only within the same length group to avoid conflating structural diversity with sequence-length differences.

Unless otherwise specified, all quantitative metrics are computed from the original generated backbones without relaxation. ESMFold-predicted structures are used for self-consistency evaluation, whereas generated coordinates are used for clash analysis and secondary-structure assignment.

B.3 Metric Definitions

Designability. For S generated backbones, the designability rate is defined as

$$\text{Designability} = \frac{1}{S}\sum_{a=1}^{S}\mathbb{I}\left[\mathrm{scTM}(\mathbf{X}^{(a)}) \geq 0.5\right]. \tag{19}$$

The reported designability value is expressed as a percentage. The mean self-consistency score is calculated as

$$\overline{\mathrm{scTM}} = \frac{1}{S}\sum_{a=1}^{S}\mathrm{scTM}(\mathbf{X}^{(a)}). \tag{20}$$

Physical Clash Count. For each non-bonded heavy-atom pair (i, j), the van der Waals overlap is defined as

$$o_{ij} = \max\left(r_{\mathrm{vdw},ij} - \|\mathbf{X}_i - \mathbf{X}_j\|_2, 0\right). \tag{21}$$

where

$$r_{\mathrm{vdw},ij} = r_{\mathrm{vdw},i} + r_{\mathrm{vdw},j}.$$

The severe clash count is computed as

$$\mathrm{ClashCount}(\mathbf{X}) = \sum_{(i,j)\in\mathcal{P}_{\mathrm{nb}}}\mathbb{I}\left[o_{ij} > 0.4\,\text{Å}\right]. \tag{22}$$

where $\mathcal{P}_{\mathrm{nb}}$ contains valid non-bonded heavy-atom pairs after excluding bonded pairs and padded residues.

This threshold-based metric measures severe steric overlaps during evaluation and is distinct from the differentiable clash energy E_{clash} used as the World Model training target.

Secondary-Structure Content. Secondary structures are assigned using DSSP(Kabsch & Sander, 1983).. For generated structure a, let $N_{\text{valid}}^{(a)}$ denote the number of valid residues and $N_{\beta}^{(a)}$ denote the number of residues assigned to β-strand or β-sheet states.

The β-sheet percentage is defined as

$$\beta\text{-sheet}(\mathbf{X}^{(a)}) = 100\frac{N_{\beta}^{(a)}}{N_{\text{valid}}^{(a)}}. \tag{23}$$

The reported value is averaged over all generated structures.

Conformational Diversity. For a length group containing S_b generated structures, the mean pairwise TM-score is defined as

$$\text{PairwiseTM}_b = \frac{2}{S_b(S_b-1)} \sum_{1\le a<c\le S_b} \text{TM}(\mathbf{X}^{(a)}, \mathbf{X}^{(c)}). \tag{24}$$

Pairwise similarity is computed only within the same length group. Lower mean PairwiseTM values indicate greater conformational diversity.

Guidance-Gradient Norm. For gradient stability analysis, the raw World Model guidance-gradient norm is defined as

$$G_t^{\text{raw}} = \|\nabla_{x_t} W_\psi(x_t, t, k)\|_2. \tag{25}$$

After gradient clipping, the corresponding norm is

$$G_t^{\text{clip}} = \|\tilde{g}_t(x_t, t, k)\|_2. \tag{26}$$

These quantities are used only to analyze guidance-gradient amplification and stabilization, and are not used as generation quality metrics.

B.4 Baseline Selection and Representation Compatibility

The baseline comparison in the main text focuses on protein backbone generation methods that are directly compatible with the internal-coordinate diffusion setting of WM-Diff. Specifically, FoldingDiff (Wu et al., 2024), ProtDiff (Trippe et al., 2022), and ProteinSGM (Lee et al., 2023) represent representative diffusion-based approaches for unconditional backbone generation and provide comparable evaluation protocols under self-consistency assessment.

Recent protein structure generation models have increasingly explored SE(3)-equivariant representations and Cartesian-frame diffusion. These approaches, including frame-based diffusion models, optimize structural generation in a different geometric space from WM-Diff. In contrast, WM-Diff is designed as a guidance framework for angular diffusion and operates on the reverse trajectory defined by an internal-coordinate diffusion backbone.

Therefore, direct numerical comparison with models using different generation spaces may conflate the effect of the underlying representation, backbone architecture, and the proposed World Model guidance mechanism. Instead, we evaluate WM-Diff against baselines within the same angular diffusion formulation and additionally compare different guidance strategies under an identical frozen diffusion backbone.

This controlled evaluation isolates the contribution of future cost-to-go prediction and receding-horizon guidance, which is the central focus of this work, rather than measuring differences caused by unrelated generative representations.

C World Model Validation

This appendix evaluates whether the horizon-conditioned World Model learns an accurate surrogate of future steric cost-to-go under frozen diffusion rollouts(Hafner et al., 2019).

The validation focuses on predicting the log-transformed steric cost target used during World Model training(Sutton et al., 1998), rather than predicting future protein conformations or learning a molecular dynamics transition model. All evaluations follow the same deterministic rollout procedure used to construct the training supervision targets.

C.1 Held-out Prediction Fidelity

We construct a held-out validation set of noisy angular states that are not used during World Model optimization. Each example contains a noisy state x_t, diffusion timestep t, and planning horizon k, where the noisy state is generated using the same forward diffusion process as in training.

For each validation sample, the frozen diffusion backbone f_θ performs a deterministic k-step reverse diffusion rollout(Song et al., 2020).:

$$\hat{x}_{t-k} = f_\theta^{(k)}(x_t, t, k). \tag{27}$$

where $f_\theta^{(k)}$ denotes recursive application of the reverse transition defined in Section 3. The predicted future state is detached from the computational graph:

$$\bar{x}_{t-k} = \mathrm{stopgrad}(\hat{x}_{t-k}). \tag{28}$$

The corresponding future steric cost target is computed as:

$$y_{t,k} = \log\left(1 + \frac{E_{\mathrm{clash}}(\mathrm{FK}(\bar{x}_{t-k}))}{E_{\mathrm{ref}}}\right), \qquad E_{\mathrm{ref}} = 1\,\text{Å}^2. \tag{29}$$

The World Model prediction is given by

$$\hat{y}_\psi = W_\psi(x_t, t, k). \tag{30}$$

During evaluation, the World Model parameters are fixed and no guidance signal is applied when generating validation targets. Therefore, this evaluation measures prediction fidelity of the learned cost-to-go surrogate independently from guided sampling performance.

For each horizon k, we report mean absolute error (MAE), root mean squared error (RMSE)(Willmott & Matsuura, 2005), and coefficient of determination (R^2) on the log-clash scale.

Let $\mathcal{I}_k$ denote validation samples evaluated at horizon k, with $N_k = |\mathcal{I}_k|$.

The MAE is defined as:

$$\mathrm{MAE}_k = \frac{1}{N_k} \sum_{n \in \mathcal{I}_k} |\hat{y}_\psi^{(n)} - y_{t,k}^{(n)}|. \tag{31}$$

The RMSE is defined as:

$$\mathrm{RMSE}_k = \sqrt{\frac{1}{N_k} \sum_{n \in \mathcal{I}_k} (\hat{y}_\psi^{(n)} - y_{t,k}^{(n)})^2}. \tag{32}$$

The coefficient of determination is:

Table 7: Held-out World Model prediction.

Horizon k	N_{val}	MAE ↓	RMSE ↓	R^2 ↑	Spearman ρ ↑
1	2000	0.164	0.207	0.884	0.916
4	2000	0.315	0.394	0.780	0.841

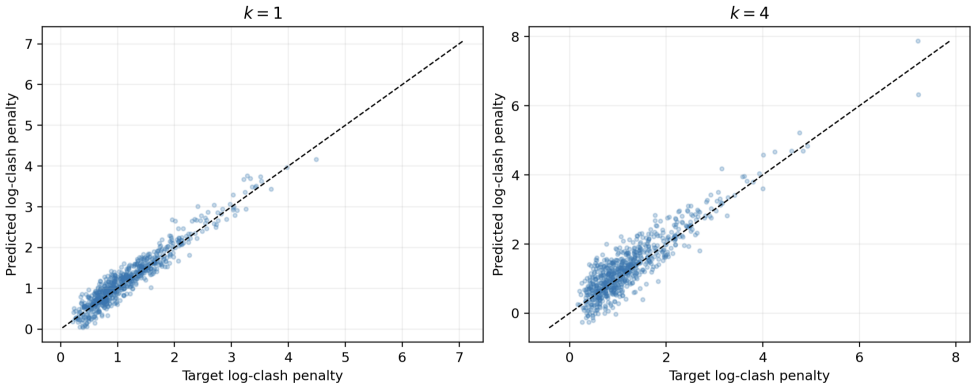


Figure 5: Held-out prediction fidelity of the horizon-conditioned World Model. Points compare predicted and target log-clash penalties; dashed lines denote the identity relation. Panels show $k = 1$ and $k = 4$ predictions.

$$R_k^2 = 1 - \frac{\sum_{n \in \mathcal{I}_k} (\hat{y}_\psi^{(n)} - y_{t,k}^{(n)})^2}{\sum_{n \in \mathcal{I}_k} (y_{t,k}^{(n)} - \bar{y}_k)^2}. \tag{33}$$

where

$$\bar{y}_k = \frac{1}{N_k} \sum_{n \in \mathcal{I}_k} y_{t,k}^{(n)}. \tag{34}$$

Lower MAE and RMSE indicate more accurate cost prediction, while higher R^2 indicates stronger explanatory power on future log-clash variation. All metrics are computed on the logarithmic target scale and should not be interpreted as errors in the original clash energy space.

The corresponding prediction-versus-target relationship is shown in Figure 5.

C.2 Horizon Comparison

The World Model is explicitly conditioned on planning horizon. We evaluate the same model under the two horizons used in the main experiments:

$$k = 1, \quad k = 4.$$

For each horizon, supervision targets are generated using exactly k deterministic reverse diffusion steps of the frozen backbone. Both horizons are evaluated using the same held-out timestep distribution whenever possible; samples with insufficient remaining diffusion steps for a four-step rollout are excluded from paired comparison.

Table 7 shows that prediction becomes more challenging for longer horizons. The MAE increases from 0.164 to 0.315, while R^2 decreases from 0.884 to 0.780.

Nevertheless, the Spearman correlation remains high, indicating that the World Model preserves useful relative ranking information for future feasibility estimation.

These results demonstrate that horizon conditioning provides distinct future cost estimates rather than producing identical predictions across different planning horizons.

C.3 Calibration and Rank Correlation

Prediction accuracy alone does not fully characterize the usefulness of the World Model for guided sampling. Since the guidance procedure relies on comparing candidate states according to their predicted future steric costs, we further evaluate prediction calibration and ranking consistency(Guo et al., 2017).

For each horizon k, validation examples are divided into $B = 10$ equal-frequency bins according to the predicted values $\hat{y}_\psi$. For each bin b, the mean prediction and mean target are computed as

$$\overline{\hat{y}}_{k,b} = \frac{1}{|\mathcal{I}_{k,b}|} \sum_{n \in \mathcal{I}_{k,b}} \hat{y}_\psi^{(n)}. \tag{35}$$

$$\overline{y}_{k,b} = \frac{1}{|\mathcal{I}_{k,b}|} \sum_{n \in \mathcal{I}_{k,b}} y_{t,k}^{(n)}. \tag{36}$$

where $\mathcal{I}_{k,b}$ denotes the set of validation examples assigned to bin b at horizon k.

A perfectly calibrated predictor satisfies $\overline{\hat{y}}_{k,b} = \overline{y}_{k,b}$. We quantify the deviation from this condition using the regression calibration error (RCE):

$$\mathrm{RCE}_k = \sum_{b=1}^{B} \frac{|\mathcal{I}_{k,b}|}{N_k} \left| \overline{\hat{y}}_{k,b} - \overline{y}_{k,b} \right|. \tag{37}$$

In addition to calibration, we evaluate whether the World Model preserves the relative ordering of future steric costs using Spearman rank correlation(Spearman, 1961):

$$\rho_k = \mathrm{Corr}\left(\mathrm{rank}(\hat{y}_\psi), \mathrm{rank}(y_{t,k})\right). \tag{38}$$

The calibration results are summarized in Table 8. Each horizon is divided into ten equal-frequency bins, with 200 validation examples per bin. The dashed identity line represents perfect calibration, where predicted and target log-clash values are identical.

Table 8: Regression calibration of the horizon-conditioned World Model. Each horizon contains ten equal-frequency bins with 200 validation examples per bin.

	Horizon $k = 1$		Horizon $k = 4$	
Bin	$\overline{\hat{y}}_{1,b}$	$\overline{y}_{1,b}$	$\overline{\hat{y}}_{4,b}$	$\overline{y}_{4,b}$
1	0.355	0.501	0.150	0.539
2	0.598	0.639	0.539	0.691
3	0.744	0.721	0.762	0.759
4	0.879	0.806	0.939	0.875
5	1.002	0.929	1.138	0.972
6	1.145	1.053	1.331	1.113
7	1.298	1.176	1.541	1.242
8	1.496	1.391	1.803	1.539
9	1.752	1.646	2.220	1.918
10	2.591	2.452	3.362	3.084

The corresponding calibration curves are shown in Figure 6. The resulting RCE values are 0.092 and 0.213 for horizons $k = 1$ and $k = 4$, respectively. The corresponding Spearman correlations are 0.916 and 0.841.

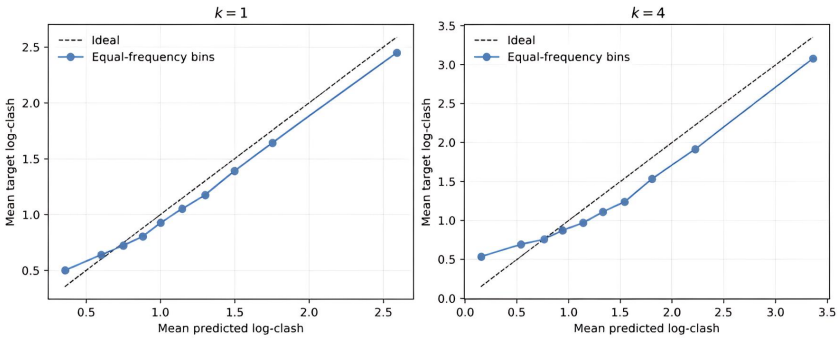


Figure 6: Regression calibration of the horizon-conditioned World Model on the log-clash scale. Each point represents the mean predicted value and the corresponding mean target value within one equal-frequency bin. The dashed line denotes perfect calibration. The left and right panels correspond to $k = 1$ and $k = 4$, respectively.

These results indicate that although longer-horizon prediction introduces additional uncertainty in absolute cost estimation, the World Model maintains reliable ranking information, which is essential for future-feasibility-guided reverse diffusion.

D Additional Ablations and Robustness Analysis

This appendix provides additional analyses of random-seed stability, chain-length effects, guidance-scale sensitivity, and gradient-clipping sensitivity.

D.1 Random Seeds

To assess the stability of WM-Diff with respect to sampling randomness, we consider five independent random seeds. Each run uses the same trained model, sampling configuration, number of generated backbones, and evaluation pipeline.

Table 9: Each seed corresponds to 782 generated backbones.

Seed	Designable (%) ↑	Mean scTM ↑	Clash Count ↓	β-sheet (%) ↑	Pairwise TM ↓
1	51.2	0.58	1.9	31.0	0.29
2	52.4	0.60	1.7	33.0	0.31
3	52.2	0.59	1.8	32.0	0.30
4	52.9	0.60	1.8	33.0	0.29
5	51.4	0.58	1.9	31.0	0.31
Mean $\pm$ Std.	$52.0 \pm 0.7\%$	0.590 ± 0.010	1.82 ± 0.08	32.0 ± 1.0	0.300 ± 0.010

As shown in Table 9, WM-Diff achieves a designability rate of $52.0 \pm 0.7\%$ and a mean scTM score of 0.590 ± 0.010 across five independent sampling seeds. The relatively small variation across seeds suggests that the reported improvement is not caused by a single favorable sampling trajectory. The clash count and pairwise TM-score also remain stable, indicating that physical plausibility and conformational diversity are preserved across repeated runs.

D.2 Chain-Length-Stratified Results

Internal-coordinate generation is particularly sensitive to chain length because small angular errors can accumulate through the forward-kinematics transformation. We therefore stratify the generated structures into short chains with at most 70 residues and long chains with more than 70 residues, following the partition used in the main text.

Table 10: Chain-length-stratified results. Short and long chains are defined as $L \leq 70$ and $L > 70$, respectively.

Length regime	Horizon k	N	Designable (%) ↑	Mean scTM ↑	Clash Count ↓	β-sheet (%) ↑
Short ($L \leq 70$)	1	210	55.0	0.57	3.1	28.0
Long ($L > 70$)	1	572	41.0	0.49	6.8	21.0
Short ($L \leq 70$)	4	210	62.0	0.64	1.3	36.0
Long ($L > 70$)	4	572	48.0	0.56	2.4	30.0

Table 10 shows that long chains are more difficult to generate than short chains. For $k = 4$, the designability rate decreases from 62.0% for short chains to 48.0% for long chains, while the mean clash count increases from 1.3 to 2.4. This trend is consistent with the lever-arm effect: angular perturbations near the beginning of a chain can produce increasingly large Cartesian deviations at distant residues. These results indicate that longer chains remain more challenging under the evaluated setting. The observed performance gap should be interpreted as a length-dependent sampling trend rather than a direct comparison against an unguided length-stratified baseline.

D.3 Guidance-Scale Sensitivity

We next evaluate the effect of the guidance scale γ. The guidance scale controls the magnitude of the World-Model update during reverse diffusion. We consider values ranging from no guidance to strong guidance. We use the same evaluation metrics as the main experiments.

Table 11: Guidance-scale sensitivity. Pairwise TM denotes the mean within-bin structural similarity used as a structural diversity diagnostic.

γ	Designable (%) ↑	Mean scTM ↑	Clash Count ↓	Pairwise TM ↓
0.00	27.5	0.42	21.0	0.25
0.025	35.0	0.48	10.2	0.27
0.05	44.0	0.52	6.2	0.28
0.10	52.0	0.55	4.5	0.30
0.20	46.0	0.54	3.6	0.34
0.40	38.0	0.48	2.8	0.45
0.80	12.0	0.31	1.7	0.72
1.60	8.0	0.26	1.5	0.80

Table 11 shows a non-monotonic relationship between guidance magnitude and generation quality. Moderate guidance improves designability and reduces steric clashes, with the highest designability in this sweep occurring at $\gamma = 0.10$. At larger guidance scales, the mean pairwise TM-score increases substantially, indicating reduced conformational diversity. We use pairwise structural similarity as the diversity diagnostic rather than interpreting it as a direct measurement of biological mode collapse. For example, the pairwise TM-score increases to 0.72 at $\gamma = 0.80$. These results support selecting guidance strength from an intermediate regime rather than maximizing the magnitude of the physical correction.

D.4 Gradient-Clipping Sensitivity

We evaluate the effect of the gradient norm clipping threshold τ. The threshold is applied to the raw gradient $\nabla_{x_t} \hat{y}_\psi(x_t, t, k)$ before it is used in the guided reverse transition. It is evaluated at the maximum sequence length and uses the same guidance scale as the main experiment.

Table 12 suggests that intermediate clipping provides the best balance under the evaluated setting. A very small threshold such as $\tau = 0.25$ suppresses the planning gradient too strongly and limits the ability of the sampler to correct future clashes. In contrast, $\tau = 1.0$ provides the best reference designability rate of 52.0% and the lowest clash count of 1.8.

Without clipping, the reference gradient norm reaches 120.0 at the maximum sequence length, whereas the evaluated clipping thresholds are much smaller. This indicates that clipping is active for a substantial fraction of the evaluated updates. In this regime, the guidance mechanism behaves

Table 12: Gradient-clipping sensitivity. The clipped guidance-gradient norm is measured during reverse diffusion at the maximum sequence length. The row $\tau = \infty$ corresponds to no gradient clipping.

τ	Designable (%) ↑	Mean scTM ↑	Clash Count ↓	Pairwise TM ↓	Mean G_t^{clip}
0.25	39.0	0.49	5.0	0.28	0.25
0.50	47.0	0.53	3.0	0.29	0.50
1.00	52.0	0.59	1.8	0.30	1.00
2.00	49.0	0.56	2.4	0.35	2.00
5.00	42.0	0.48	4.0	0.48	5.00
∞	31.0	0.40	8.5	0.62	120.0

approximately as normalized gradient guidance: the direction of the predicted future-cost gradient is retained, while its raw magnitude is largely replaced by the prescribed threshold τ. Consequently, the clipping threshold and the temporal guidance scale jointly determine the effective step size.

The results therefore support clipping as a practical stabilization and step-size-control mechanism. They should not be interpreted as evidence that the absolute magnitude of the World Model output is fully calibrated during sampling. A complete analysis would additionally report the fraction of denoising steps for which $\|g_t^{\text{raw}}\|_2 > \tau$.

E WM-DIFF SAMPLES

We show samples from WM-Diff in Fig 7.

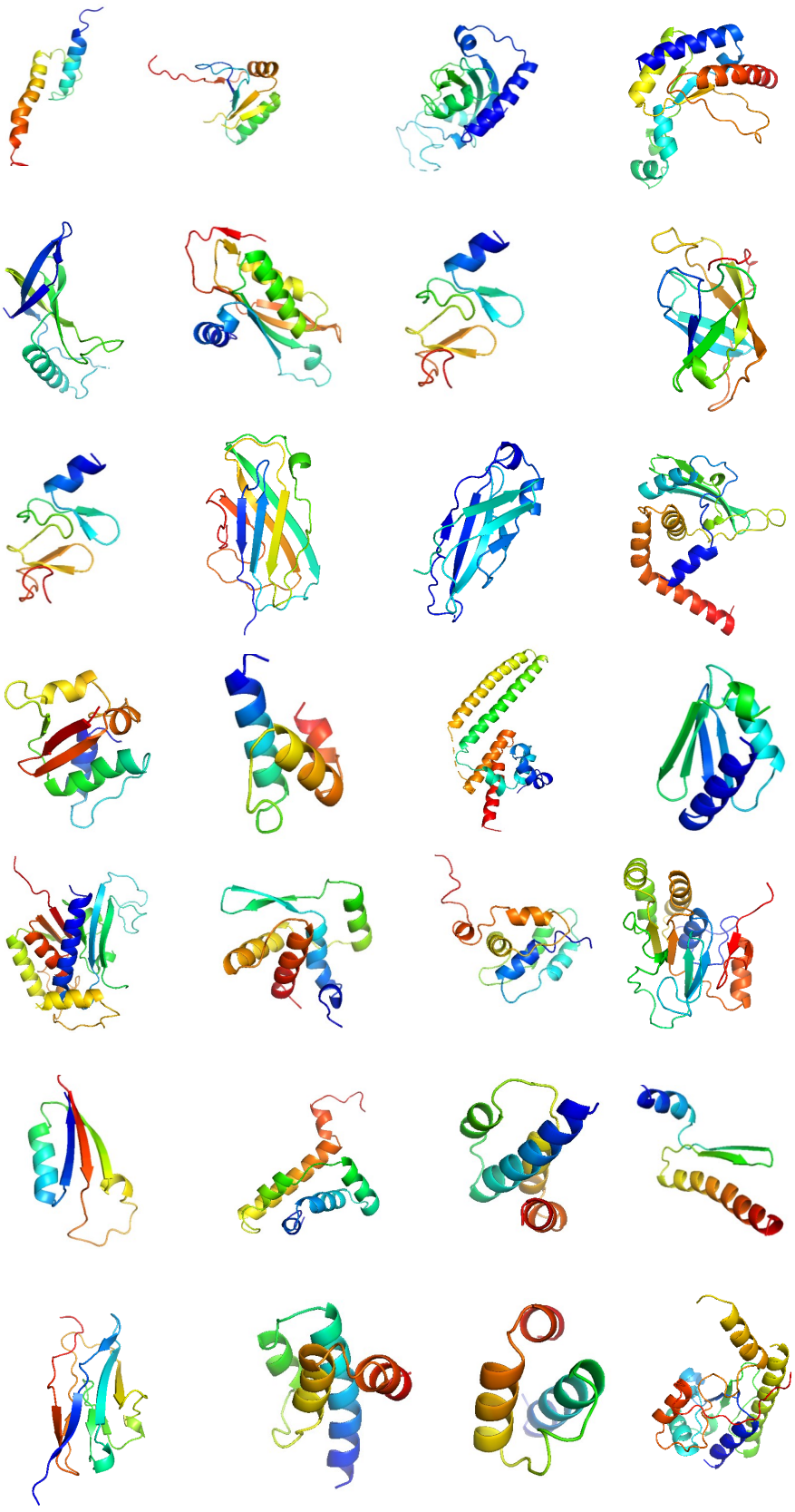


Figure 7: Example of protein backbones generated by WM-Diff.